



Online-Publikationen des Arbeitskreis Studium Populärer Musik e.V. (ASPM)

Hg. v. Ralf von Appen, André Doehring u. Thomas Phleps

www.aspm-samples.de/samplesii/Samples11/fleischer.pdf

Jahrgang 11 (2012) - Version vom 21.11.2012

DIE MÖGLICHKEITEN UND GRENZEN DER MESSBARKEIT VON MUSIKGESCHMACK. EINE KRITISCHE AUSEINANDERSETZUNG MIT DER STUDIENREIHE »THE DO RE MI'S OF EVERYDAY LIFE: THE STRUCTURE AND PERSONALITY CORRELATES OF MUSIC PREFERENCES«

Benedikt Fleischer

Einleitung

Das Konstrukt Musikgeschmack in der Zukunft zu erfassen und messbar zu machen, kann als ein die Disziplinen verbindendes Ziel aufgefasst werden, das ihnen insgesamt nutzt, um den Menschen besser zu begreifen und sein zukünftiges Verhalten genauer vorhersagen zu können. Die Musikgeschmacksforschung ist auch deshalb ein Feld mit großem Forschungspotential, weil sie eine hohe Relevanz und Bezug zum Alltagsleben des Menschen aufweist. Dies erkannten auch die Psychologen Peter Rentfrow und Samuel Gosling und veröffentlichten 2003 ihre Studienreihe »The Do Re Mi's of Everyday Life: The Structure and Personality Correlates of Music Preferences«, in der sie nicht nur einen Test zur Musikgeschmacksmessung, den *Short Test of Musical Preferences* (Rentfrow/Gosling 2003b), entwickelten, sondern auch den Zusammenhang zwischen Persönlichkeit und Musikgeschmack sowie verschiedene Dimensionen des Musikgeschmacks genauer untersuchten.

Die Studienreihe erlangte international Aufmerksamkeit (Behne 2007: 435-436). Es handelt sich zwar nicht um die erste Studie, die einen Geschmackstest erstellt hat (vgl. Cattell/Anderson 1954; Litle/Zuckerman 1986), doch eignet sie sich hinsichtlich Forschungsumfang bzw. -breite und

die Verwendung von relativ aktuellen Daten besonders gut für eine kritische Auseinandersetzung mit den Chancen und Grenzen der Messbarkeit von Musikgeschmack. Diese Kritik gewährleistet, dass die Möglichkeiten und die Theorien der Musikgeschmackforschung voll ausgeschöpft werden und gleichzeitig vorhandene Probleme, nicht beachtete Aspekte und falsche Annahmen aufgedeckt und in der weiteren Forschung berücksichtigt werden können. Nur so kann es langfristig gelingen, ein objektives, reliables und valides Messinstrument für Musikgeschmack zu entwickeln. Da Musikgeschmack sowohl psychologische, soziologische als auch kulturwissenschaftliche Forschungsfelder berührt, ist das interdisziplinäre Arbeiten Voraussetzung für den Erfolg dieser Forschung.

Dieser Artikel soll am Beispiel STOMP die wichtigsten Kritikpunkte an Testinstrumenten, die Musikgeschmack und -präferenzen auf die gleiche Weise operationalisiert haben, zusammenfassen und Vorschläge für einen neuen Forschungsansatz bieten. Die zentrale Fragestellung lautet dabei:

Sind verbale Musikpräferenztests wie der STOMP, bei denen Items in Form von Musikgenres auf Ratingskalen eingestuft werden, vor dem Hintergrund der bisherigen Musikgeschmacksforschung und Testtheorie ein valides Instrument zur Messung des Konstruktes Musikgeschmack?

Methode und Vorgehensweise

Die Studienreihe von Rentfrow und Gosling ist in sechs Teilstudien gegliedert. In der ersten Studie untersuchen sie die Bedeutung von Musik und Musikpräferenzen für Menschen im Alltag, woraufhin sie in der zweiten Studie den STOMP entwickeln und eine Faktorenanalyse auf Basis der mit diesem Verfahren erhobenen Daten durchführen, um Muster hinter erfassten Musikpräferenzen auszumachen. Die darauffolgenden beiden Studien setzen sich mit der Generalisierung der Ergebnisse über verschiedene Stichproben, Erhebungsmethoden und geographische Regionen hinweg auseinander, während in Studie 5 versucht wird die ermittelten Geschmacksdimensionen genauer zu attribuieren. Zuletzt wird der Zusammenhang zwischen Musikpräferenzen und Persönlichkeitseigenschaften, erhoben mit dem Big Five Persönlichkeitsinventar, untersucht. Der Schwerpunkt der Auseinandersetzung in diesem Aufsatz liegt auf Studie 2 der Studienreihe, denn hier wird die Testerstellung des STOMP beschrieben und gleichzeitig die Ermittlung der Musikdimensionen mithilfe des Testverfahrens erläutert. Die Studie 2 liefert somit die Basis für die gesamte Studienreihe. Schwächen in diesem Teil der Studienreihe würden sich negativ auf die nachfolgenden Studien auswirken,

weshalb eine Fokussierung auf diesen Teil sinnvoll erschien. Ebenfalls mitbezogen wurden Einleitung, Diskussion und Schlussfolgerung, sofern sie inhaltlich Bezug zur Thematik von Studie 2 nahmen. Rentfrow und Gosling erläutern in diesen Abschnitten ihre Theorien, legen ihre Argumentationsstruktur für die verwendeten Methoden dar und betten ihre Arbeit in den bisherigen Forschungskontext ein. In diesem Aufsatz erfolgt sowohl ein Abgleich der Inhalte und Aussagen der Studienreihe mit den bisherigen Forschungserkenntnissen über Musikgeschmack und Genretheorie als auch eine Untersuchung der verwendeten Forschungsmethode. Dabei werden sowohl testtheoretische Aspekte und psychologische Erkenntnisse über Musikgeschmack als auch die Literatur der Kulturwissenschaften und der Soziologie berücksichtigt, um zu gewährleisten, dass das Konstrukt in seiner Komplexität dargestellt und berücksichtigt wird.

Forschungsstand und Konstruktdefinition

Die Untersuchung von »music preferences« soll als ein Wegweiser zu einer allumfassenden Theorie fungieren, die erklären kann, »when, where, how and why people listen to music« (Rentfrow/Gosling 2003c: 1236). Rentfrow und Gosling behaupten, dass das Thema Musik, trotz der hohen sozialen Relevanz und des hohen Einflusses im Alltag, einen zu geringen Stellenwert in der Forschung eingenommen habe (2003: 1236). Zwar ist nicht zu leugnen, dass der Thematik nach dieser Studienreihe eine wesentlich höhere Bedeutung zugemessen wurde (s. z.B. Chamorro-Premuzic et al. 2009; Delsing et al. 2008; Mulder et al. 2010; Zweigenhaft 2008); doch entspricht die Darstellung als bisher vernachlässigtem Bereich in der Sozial- und Persönlichkeitspsychologie nicht den Tatsachen. Dies wird schon bei einem Blick in die Quellenangaben der Studie deutlich: Wichtige europäische Forscher wie Klaus-Ernst Behne, Rainer Dollase, Andreas Gebesmair, Ekkehard Jost, Helga de la Motte-Haber oder Helmut Rösing werden dort nicht erwähnt, obwohl sie schon seit 1965 zum Thema geforscht haben.¹

1 Dies könnte auch ein Hinweis darauf sein, dass auf internationaler Ebene in diesem Bereich noch zu wenig Austausch über die aktuelle Forschung stattfindet. So könnten stark divergierende Wissensstände daher rühren, dass zu wenig deutsche Studien in englischer Sprache veröffentlicht werden. Dies steht auch im Zusammenhang mit den Standards wissenschaftlichen Arbeitens: Die APA Standards untersagen eine doppelte Publizierung einer Studie in zwei verschiedenen Sprachen aus Gründen der Ressourcenschonung, Copyrightverletzung und Übersichtlichkeit über das wissenschaftliche Forschungsfeld (American Psychological Association Staff 2009: 13). Die Deutsche Eliteakademie beklagt diesen Zustand und verlangt in diesem Zusammenhang eine Änderung der Publikations-

Unabhängig davon, aus welchen Gründen die genannten Forscher nicht berücksichtigt worden sind, zeigt sich die Notwendigkeit der genaueren Betrachtung ihrer Schriften bei der Definition des Konstruktes Musikgeschmack sehr deutlich. Diese ist der erste Schritt für eine theoretische Basis und absolut unumgänglich zur klaren Abgrenzung des Konstruktes (Bühner 2006: 105). Bei Rentfrow und Gosling wird zu Beginn keine Definition von Musikpräferenzen dargelegt. Sie nehmen auch keine Abgrenzung zum Begriff Musikgeschmack vor. Dies scheint sich zunächst zu erübrigen, da bei der Erstellung des STOMP keine rationale Testkonstruktion, also eine Ableitung eines Testinstrumentes aus einer vorher entwickelten Theorie vorliegt. Stattdessen bevorzugen beide eine induktive Testkonstruktion. Hierbei wird eine gewisse Menge Items – in diesem Fall Musikgenres –, von denen man ausgeht, dass sie ein gemeinsames Konstrukt erfassen, mittels exploratorischer Faktorenanalyse ausgewertet, um »Dimensionen zu finden, die den Itemantworten zugrunde liegen« (Bühner 2006: 94). Eine solche Datenanalyse wird in Studie 2 vorgenommen, allerdings ohne das Konzept von Musikgeschmack und seinen Zusammenhang zu den verwendeten Genres zu erörtern (Rentfrow/Gosling 2003c: 1241). Auch im Diskussionsteil nach der Datenanalyse gehen die Forscher nicht näher darauf ein. Zwar erläutern sie ausführlich, was eine Theorie der Musikpräferenzen leisten und beantworten können müsste oder welche anderen Faktoren auf sie einwirken, aber eine abschließende Konstruktdefinition fehlt dennoch. Der Zusammenhang zwischen Musikpräferenzen und Musikgenres wird aufgrund der hohen Augenscheinvalidität der Genrebegriffe angenommen, ohne den Begriff Genre und seine Eignung als Item genauer zu betrachten. Laut Markus Bühner stellt aber die Validität, die nur auf logischen und fachlichen Überlegungen fußt, kein ausreichendes Kriterium für die Güte eines Testes dar (Bühner 2006: 62; siehe auch Lienert/Raatz 1998: 36f.).

Dabei hat die bisherige Forschung bereits eine Vielzahl unterschiedlicher Ansätze hervorgebracht, deren ausführliche Darstellung den hier gesetzten Rahmen überschreiten würde. Es ist allerdings auffällig, dass über eine einheitliche Definition von Musikgeschmack bisher kein Konsens besteht und die gesetzten Schwerpunkte auch über die Disziplinen hinweg stark vonei-

ethik, um einen weltweiten Zugriff auf deutsche Studien zu ermöglichen und sie gleichzeitig auf internationaler Ebene präsentieren zu können (2011: Abs. 5). Interdisziplinäres Arbeiten, besonders über Kontinente und Länder hinweg, könnte so vereinfacht werden und würde im Nachhinein sogar Ressourcen sparen. Ein weltweiter Zugriff auf in mehreren Sprachen publizierte Studien könnte den wissenschaftlichen Austausch wesentlich erleichtern und helfen die Forschungsstände in allen Ländern auf ein gleiches Niveau zu setzen, ohne dafür den persönlichen Kontakt notwendig zu machen.

nander divergieren (vgl. Niketta 1993: 339; Dollase 1998: 342f.; Jost 1982: 246f.; Lehmann 1994: 111; Gebesmair 2001: 48ff.; Behne 1986: 19).

Schon die Anzahl theoretischer Ansätze zeigt, dass zahlreiche unterschiedliche Theorien im Bezug auf Musikgeschmack und -präferenzen existieren, auf deren Basis eine rationale Testkonstruktion möglich gewesen wäre, und dass das Forschungsgebiet somit nicht mehr das Neuland darstellt, welches man zunächst mit induktiven Methoden zu beschreiben versuchen würde.

Der Short Test of Music Preferences und seine Genrekategorien

In ihrer zweiten Studie gehen Rentfrow und Gosling (2003c: 1240ff.) auf die Theorie und Erstellung des STOMP ein und analysieren die mit ihm erhobenen Daten von 1704 Studenten der University of Texas in Austin mithilfe einer explorativen Faktorenanalyse. Anschließend interpretieren sie die gefundenen Faktoren und ermitteln die Retestreliabilität.

Der STOMP kann sich nicht eindeutig in die Klassifizierung von Tests nach Gustav A. Lienert und Ulrich Raatz (1998: 15) eingeordnet werden, was mit der unklaren Konstruktdefinition zusammenhängt. Zumindest lässt er sich als ein psychometrisches, verbales Individualtestverfahren definieren, bei dem 14 Genres (Classical, Blues, Country, Dance/Electronica, Folk, Rap/Hip-Hop, Soul/Funk, Religious, Alternative, Jazz, Rock, Pop, Heavy Metal, Soundtracks/theme songs) vom Teilnehmer einzeln auf einer bipolaren 7-stufigen Bewertungsskala von »Dislike Strongly« bis »Like Strongly« eingestuft werden. Viele der verwendeten Items erweisen sich bei genauerer Betrachtung bereits als ungeeignet, da sie entweder aus zwei Genres zusammengesetzt wurden, die verschiedene historische Entwicklungen haben und unterschiedlich präferiert werden könnten (z.B. Soul/Funk, vgl. Wicke/Ziegenrucker/Ziegenrucker 2007: 268 u. 680) oder in sich so heterogen und breit gefächert sind, dass sie sämtliche anderen Items des Tests beinhalten könnten, wie es bei Soundtrack/theme songs der Fall ist. Dadurch wird unklar, an welche Art der Musik die Versuchsperson beim Evaluieren denkt. Nicola Sigg (2009: 24) merkt außerdem an, dass die Genres nicht in alle Kulturen übertragbar seien, da ihre Verbreitung je nach Region variieren könne und der Test immer entsprechend angepasst werden müsste. Diese Problematik wird im STOMP-Revised, den Rentfrow und Gosling im Internet zur Verfügung stellen, jedoch ebenfalls nicht berücksichtigt und es erfolgt keine entsprechende Verbesserung der Items. Die Zahl der zu bewertenden Genres

wird von 14 auf 23 aufgestockt (hinzukommen Bluegrass, Gospel, International/Foreign, New Age, Oldies, Opera, Punk und Reggae) und die Kombination Soul/Funk wird aufgeteilt in Soul/ R&B und Funk. Allerdings wird keine dieser Veränderungen als Optimierung des Testverfahrens begründet. Es ist also nicht ersichtlich, warum der STOMP in diesen Punkten revidiert wurde (Rentfrow/Gosling 2011).

Rentfrow und Gosling berufen sich in ihrer Studie unter anderem auf John/Hampson/Goldberg (1991) und argumentieren, dass der optimale und am häufigsten gebrauchte Abstraktionsgrad für Musikpräferenzen die Genre- und Subgenreebenen seien (Rentfrow/Gosling 2003c: 1241). Diese Ebenen wären gleichzusetzen mit dem sogenannten »Basic Level«, der Kategorie, die bei möglichst hoher Breite noch den höchstmöglichen Informationsgehalt über einzelne eingeordnete Objekte besitzt (John/Hampson/Goldberg 1991: 348). Solche Eigenschaften treffen aber auf die Genres Soundtrack/theme song und International/Foreign nicht zu, da sie so breit sind, dass sie alle anderen Genres des Tests beinhalten können. Darüber hinaus lässt sich durch reine Betrachtung nicht bestimmen, dass die Kategorie Genre wirklich das Basic Level im Bereich Kategorisierung von Musik darstellt. John/Hampson/Goldberg (ebd.: 350f.) führten für ihren Themenbereich »Wertekonzepte« erst eine eigene Studie durch, um das Basic Level bei Adjektiven zu bestimmen, die menschliche Eigenschaften beschreiben. Ob dies auf die Beschreibung von Musik übertragbar ist, ist nicht gesichert. Die Ermittlung der besten Items nach Vertrautheit, wie bei Rentfrow und Gosling (2003c: 1241) geschehen, erscheint sinnvoll, da man nur das Musikgenre bewerten kann, das man kennt. Es ist aber falsch zu glauben, damit garantiert die Basic Level-Kategorie zu ermitteln, da John/Hampson/Goldberg (1991: 357) in derselben von Rentfrow und Gosling zitierten Studie nachweisen konnten, dass zwischen der Vertrautheit mit einem Thema und der Verwendung von Basic Level Kategorien kein unmittelbarer Zusammenhang besteht. Sie ermöglicht zwar das Abrücken von den noch breiter gefassten »Superordinate« Level Kategorien, aber garantiert nicht die Verwendung von Basic Level Kategorien. Dies würde auch die teilweise zu allgemein gefassten Items erklären.

In der Studienreihe von Rentfrow und Gosling sind zudem testtheoretische Mängel feststellbar. Dies wird bei der Betrachtung der im Internet von Gosling veröffentlichten repräsentativen Normstichprobe für den STOMP deutlich (Rentfrow/Gosling 2003a). Hier ist es bereits fragwürdig, ob ein einzelnes Testitem pro Genre ausreichend ist, um eine zuverlässige Aussage über die Präferenzen gegenüber einem Genre zu treffen, da die Reliabilität erst mit längeren Tests ansteigt (Lienert/Ratz 1998: 209). Die Tabelle der

Normstichprobe zeigt darüber hinaus die Mittelwerte und Standardabweichungen der Evaluationen für die einzelnen ermittelten Musikdimensionen an. Des Weiteren sind zwar die Daten nach Geschlecht und Ethnizitäten eingeteilt, es fehlen aber wesentliche Informationen über die Ziehung und Eichung der Stichprobe. Man weiß beispielsweise nicht, ob es sich hier um eine Gebiets- oder Quotenstichprobe handelt. Es bleibt daher unklar, wo, wann und nach welchen Kriterien sie erhoben wurde oder welchen Geltungsbereich sie umfasst (ebd.: 273f.). Somit ist unersichtlich, für welche Altersspanne und auf welche Populationen sie überhaupt anwendbar ist. Es müssten außerdem wesentlich mehr Daten zur Häufigkeitsverteilung der Testrohwerte und zur Homogenität und Varianzanalyse der gesamten Stichprobe und einzelner Gruppen geliefert werden, um einen Überblick über ihre Struktur zu erhalten (ebd.: 276ff.). Rentfrow und Gosling geben beim STOMP und auch beim STOMP-R generell zu wenig notwendige Informationen. Weder gibt es ein Testhandbuch noch Bedingungen, unter denen der Test bearbeitet werden soll, oder Angaben darüber, für welche Personengruppen er geeignet ist. Es fehlt an Durchführungs- und Auswertungsanweisungen sowie an Interpretationsbeispielen. Dabei sind dies wichtige Punkte, die für ein gutes Testverfahren absolut notwendig sind (Bühner 2006: 79).

Im Bezug auf die Testerstellung ist die Aufgabenanalyse bei der Testvorform unzureichend durchgeführt worden und weist erhebliche Lücken auf. Den 30 Testpersonen wurden 80 Genres und Subgenres als Gesamtdarbietung vorgesetzt, die aus einer freien Assoziationsaufgabe ermittelt und mit Begriffen von Online-Musikanbietern ergänzt worden sind. Diese sollten bewertet oder leer gelassen werden, falls sie einem Teilnehmer nicht bekannt waren (Rentfrow/Gosling 2003c: 1241). Im Nachhinein werden allerdings keine Informationen über die 30 ausgewählten Teilnehmer am Vortest bekannt gegeben, womit unklar ist, ob sich die Analysestichprobe vom späteren Testkollektiv unterscheidet. Darüber hinaus ist die Größe der Analysestichprobe zu gering. Sie wird bei Lienert/Raatz (1998: 60) mit mindestens 200 bis höchstens 400 Personen veranschlagt, je nach Homogenität der Grundgesamtheit. Eine anschließende Berechnung von Rohwerten, Popularitätsindizes, Trennschärfekoeffizienten und Aufgabeninterkorrelationen wird ebenfalls nicht erwähnt, obwohl solche Berechnungen sehr aufschlussreiche Informationen über die Qualität der Items hätten offenbaren können. Eine ausführliche Beschreibung der einzelnen Faktoren würde hier zwar den Rahmen sprengen, es kann aber mit Sicherheit gesagt werden, dass auf Basis eines umfassenden Vortests die bereits erwähnten schlecht definierten Items hätten identifiziert werden können. Eine Aufgabenanalyse hätte auch die Möglichkeit eröffnet, Reliabilitäts- und Validitätskoeffizienten zu er-

rechnen, um einen Überblick über die Leistungsfähigkeit des Vortests zu erhalten. Gleichzeitig wäre eine Schätzung der Gütekriterien für das Testverfahren vor und nach der Aufgabenselektion ermöglicht worden (ebd.: 134f.). Eine solche Selektion findet bei Rentfrow und Gosling nur auf der Basis des Bekanntheitsgrades der Genres statt.

Die Reliabilität des STOMP

Um eine zeitliche Stabilität der ermittelten vier Faktoren garantieren zu können, wurde der STOMP von 118 Teilnehmern aus der Stichprobe nach ungefähr drei Wochen ein weiteres Mal ausgefüllt. Anschließend wurden die ermittelten Faktoren operationalisiert und die Korrelationen zwischen den beiden Messzeitpunkten für jeden Faktor ermittelt, die ein zufrieden stellendes Ergebnis für die Retestreliabilität ergaben (Rentfrow/Gosling 2003c: 1242). Wie jedoch bereits bei der Stichprobe des Vortests ist auch hier nach den Maßen von Lienert/Raatz (1998: 174) die Stichprobengröße dieser Teilstichprobe zu klein. Sie empfehlen eine Größe zwischen 200 und 500 Teilnehmern, die aber vor allem dann größer sein sollte, wenn die Retestreliabilität als einzige Methode der Reliabilitätskontrolle verwendet wird. Auch der Rückschluss auf eine langfristige Stabilität des Persönlichkeitsmerkmals ist bei einem so kurzen Zeitraum zwischen den Erhebungen problematisch. An diesem Punkt ist es daher sinnvoll sich mit dem Geltungsbereich des Testes auseinanderzusetzen. Dazu sollte außerdem das Alter der Teilnehmer als einer der wichtigsten Einflussfaktoren auf den Musikgeschmack genauer betrachtet werden. Dollase entwickelte dazu ein entwicklungspsychologisches Modell, das die Ausbildung musikalischer Präferenzen in drei Phasen teilt (Dollase 1998: 356).² In der zweiten Studie von

2 In der Anstiegsphase, die ungefähr ab dem 10. Lebensjahr beginnt, nimmt die Orientierung an den Präferenzen der Eltern langsam ab. Gleichzeitig nimmt der Einfluss von Gleichaltrigen, welcher sich in sehr gruppenkonformen Präferenzen und eingeschränkter Offenheit äußert, zu. Die Dauer des Musikhörens steigt an, genau wie die zunehmende Emotionalisierung der Musik. Jugendliches Musikinteresse gipfelt daraufhin in der Plateauphase. Mit ca. 13 Jahren verbindet sich das Musikinteresse mit anderen kulturellen und sozialen Bereichen, politischen Orientierungen und Lebensweisen. Der Musikkonsum wird sozial und kognitiv überformt, es bilden sich individuelle und soziale Funktionen aus, die in den Alltag integriert werden. Musik wird Informationsquelle und Erkennungszeichen für Moden, Lebensstile und Verhaltensweisen, fungiert als Stimulanz für Träume und Sehnsüchte und wird zur Bewältigung bestimmter Zustände und Situationen verwendet. Gleichzeitig differenzieren sich Hörweisen und -gewohnheiten zunehmend aus, wodurch zunehmend von den Peers unabhängige Interessenprofile und Präferenzen entstehen, was mit der Identitätskonstruktion in

Rentfrow/Gosling gibt es keine Angaben zum Mittelwert noch zur Standardabweichung in Bezug auf die Altersverteilung der Stichprobe. Allerdings lässt die Stichprobe der ersten Studie mit einem Durchschnittsalter von 18,9 Jahren und einer Standardabweichung von 2,3 den Rückschluss auf eine vergleichbare Altersspanne zu, da die Population vom Studiengang derselben Universität im Jahre 2001 stammt (Rentfrow/Gosling 2003c: 1238). Damit liegt die Stichprobe der zweiten Studie genau in der von Dollase beschriebenen instabilen Phase der musikalischen Sozialisation. Das bedeutet für die Retestrelabilität, dass sie nur eine eingeschränkte Aussagekraft für die langfristige Stabilität der Faktoren über die drei Wochen hinaus besitzt. Denn der häufige und schnelle Wechsel von Präferenzen als Ausdruck der Suche nach einer eigenen Identität wird zusätzlich durch die rasche Abfolge von Trends und Idolen forciert, sodass schwer einschätzbar ist, wie stabil die ermittelten Faktoren nach einigen Monaten oder einem Jahr noch sind. Auch Behne stützt diese Entdeckung, wenn er die Altersspanne zwischen 10 und 20 Jahren als jene mit den stärksten Veränderungen kennzeichnet (Behne 1986: 55f.). Damit stellt diese Studie lediglich die Präferenzdimensionen dieser Population von Studenten in einem kurzen Zeitraum dar. Zusätzlich wird nicht berücksichtigt, dass die Teilnehmer unterschiedlich weit in der Entwicklung ihrer Präferenzen sind. Damit ist diese Altersgruppe zur Ermittlung stabiler Dimensionen nicht geeignet.

Die Validität des STOMP

Die Betrachtung des Gütekriteriums der Validität ist unmittelbar mit der zentralen Fragestellung der Arbeit verbunden. Nach der Definition des Messgegenstandes erfolgt normalerweise die Festlegung eines Validitäts-

der Pubertät einhergeht. Musikgeschmack wird Teil einer selbst bestimmbaren Identität und Persönlichkeit. Dieser Prozess setzt sich bis ca. Mitte 20 fort. Die darauf folgende Abschwungphase ist gekennzeichnet durch eine Abnahme des Musikinteresses und -hörens sowie des emotionalen Engagements in Musik, bedingt durch die Zunahme von beruflichen und familiären Interessen. Musik und ihre Funktionalisierung nehmen einen zunehmend geringeren Stellenwert ein, u.a. da andere Ziele nun mehr Vorrang haben, weniger Zeit für den Musikkonsum vorhanden ist oder auch die in der Jugend erfolgreiche Funktionalisierung von Musik nun den Sachzwängen des Erwachsenenlebens nicht mehr gewachsen ist (Dollase 1998: 360). Dollase fügt bei seinem Modell an, dass die Altersangaben flexibel zu handhaben seien, da die Entwicklung jedes Individuums unterschiedlich verläuft und fließende Übergänge zwischen den Phasen der Realität mehr entsprechen als klare Grenzen. Insbesondere die Abschwungphase könne durch eine verlängerte Jugendphase, die Postadoleszenz, hinausgezögert werden (ebd.: 360f.).

bereiches. Es muss dabei bestimmt werden, welcher Bereich des Merkmals Musikpräferenz vom Test umschlossen werden soll. Konkret bedeutet dies ein Kriterium zu finden, welches wiederum diesen Bereich widerspiegelt und an dem eine umfassende Validierung möglich ist (Lienert/Raatz 1998: 40f.). Da der Test sich bei seinen Items auf der Ebene der Genres bewegt, müsste man ein ebenso breit gefasstes Kriterium zur Validierung finden. Rentfrow und Gosling haben jedoch die Validierung ihres Tests vollkommen übergangen. Es gibt keine Angaben für eine empirische bzw. kriterienbezogene Validität, die zeigen könnten, dass der Test mit einem Außenkriterium korreliert. Es ist darüber hinaus unklar, ob der Test über eine gewisse Vorhersagekraft, auch prognostische Validität genannt, verfügt und die Vorhersage von Kriterienwerten ermöglicht (ebd.: 223).

Die bereits erwähnte Instabilität der Präferenzen in dieser Altersgruppe schränkt eine Vorhersagekraft des Tests zusätzlich noch einmal ein. Es findet darüber hinaus keine Korrelation mit anderen Tests statt, die ebenfalls Musikpräferenzen erfassen, wobei es sich schwierig gestalten dürfte, einen solchen bestätigten validen Test zu finden, der zur Überprüfung der inneren Validität verwendet werden kann. Andere Testverfahren wie der »Music Preference Questionnaire« (MPQ), ursprünglich von Paul Sikkema entwickelt, oder der »Music Preference Scale« (MPS) von Patrick Litle und Marvin Zuckerman sind nach den gleichen Prinzipien wie der STOMP erstellt worden und damit ebenso infrage zu stellen (Testverfahren verwendet in Delsing et al. 2008; Litle/Zuckerman 1986; Selfhout et al. 2007).

Die psychologische Bedeutung steht bei der kriterienorientierten Validität nur im Hintergrund. Sie ist Schwerpunkt bei der Analyse der Konstruktvalidität. Um einen möglichst umfassenden Blick auf das zu messende Konstrukt Musikpräferenz zu bekommen, wäre vor der Testerstellung eine Merkmalsanalyse sinnvoll gewesen, die es ermöglicht hätte, die Komplexität des Messgegenstandes zu überblicken. Diese psychologische Analyse dient der Bestandsaufnahme, mit der man Inhalte, Eigenschaften und Verhaltensweisen sammelt, die es in diesem untersuchten Merkmalsbereich gibt. Eine solche Auseinandersetzung wäre notwendig gewesen, um das Konstrukt genauer zu erfassen und auch mögliche Teilbereiche des Merkmals voneinander abzugrenzen (Lienert/Raatz 1998: 43f.). Rentfrow und Gosling entscheiden sich jedoch sehr schnell für das Genre als Testitem, ohne die genau Beziehung zum Konstrukt zu überprüfen. Sie unterstellen dem STOMP diese inhaltliche Validität, ohne Experten der Musikgeschmacksforschung oder professionelle Musiker, die tagtäglich musikalische Urteile treffen müssen und daher mehr Erfahrung mit dem Konstrukt haben, hinzuzuziehen. Dies kann zu einer Übergeneralisierung führen, bei der die Interpretation der Er-

gebnisse über die erfassten Leistungen hinausgeht, obwohl sich die Bewertung von Genrebegriffen als unzureichend für die Erfassung von Präferenzen herausstellen kann.

Allein mit diesen Punkten wird deutlich, dass viele bereits genannte fehlende Arbeitsschritte der Testerstellung des STOMP wichtige Voraussetzungen für eine hohe Konstruktvalidität sind, welche durch die identifizierten Schwächen zusätzlich noch vermindert wird. Rentfrow und Gosling haben in dieser Studienreihe den Schwachpunkt der Konstruktvalidität, das Fehlen eines eindeutigen Maßes, als Raum für ihre Spekulation genutzt.

Musikgenres und ihre Regeln als Taxonomien des Individuums und der Gesellschaft

Franco Fabbri sieht Genres als Kategorien, die Ordnung in die schwer überschaubare Umwelt bringen sollen und dabei Objekte oder Events nach bestimmten Kriterien gruppieren, um so ihre Komplexität zu reduzieren. Voraussetzung dafür ist die Ausbildung eines sogenannten »cognitive types«, ein »private set of instructions which allows us to recognize a specific perceptual experience as an occurrence of a particular type« (Fabbri 1999: 2). Diese innere Repräsentation von Instruktionen und Beschreibungen kann im sogenannten »socialized nuclear content« visualisiert werden, indem das Individuum seine innere Repräsentation mit anderen teilt. Dies kann zu einem Aushandeln einer gemeinsamen intersubjektiven Repräsentation führen, die sich dann in einem »cultural object« wie beispielsweise einem bestimmten Musikgenre wie »Rock« zeigt. Ein solches Kulturobjekt kann wiederum Einfluss auf die Ausbildung von cognitive types haben, wenn diese durch Erziehung oder Erfahrung an den nuclear content adaptiert werden. Damit stehen nuclear contents bzw. cultural objects ständig unter gesellschaftlichem Einfluss und müssen letztendlich mit durchlässigen Grenzen ausgestattet sein, da sie in Form und Inhalt auch ständigen Veränderungen unterworfen sind. Fabbri (1999: 3) nennt als weiteren großen Einflussfaktor neben der Erziehung die Kunst, da sie alternative Schemata der Wahrnehmung anbiete, mit denen Menschen ihre Umwelt begreifen können. Damit ist auch die Musik selbst für die Ausgestaltung ihrer eigenen Kategorisierung mitverantwortlich. Ein Begriff wie Rock kann aus zahlreichen verschiedenen Perspektiven betrachtet werden und es wird daher schwer sein, die Komplexität der dahinter stehenden kulturellen Objekte durch eine Evaluation eines einzelnen Begriffes wie im STOMP zu fassen, geschweige denn einordnen zu können, aus welcher Perspektive der Begriff nun bewertet wurde. Des Wei-

teren hat der Unterschied zwischen dem cognitive type und dem socialized nuclear content zur Folge, dass beim Auswerten des STOMP unklar ist, ob eine Person einen Genrebegriff nun auf Basis ihres eigenen cognitive types bewertet hat oder auf dem socialized nuclear content. Sie könnte die bekannte Definition eines bestimmten Genres ablehnen oder einfach nicht mit ihr vertraut sein und auf dieser Basis dem Genre andere Musik oder auch einen anderen Künstler zuordnen als gesellschaftlich ausgehandelt wurde. Daraus resultiert, dass die Bewertungsgrundlage für jeden unterschiedlich ist, da die kognitiven Voraussetzungen in Bezug auf die einzelnen Genres unterschiedlich sind.

Fabbri's Aufzählung unterschiedlicher Arten von Taxonomien zeigt, dass es bei Genres für den Einzelnen nicht nur um Wissen geht, sondern auch um Überzeugungen und Vorstellungen und damit auch um Unsicherheiten und Ambivalenzen bei der Strukturierung von Musik in Genres: »Genres are more about beliefs and practice than about theory« (Fabbri 1999: 11). Diese Feststellung schränkt ihre Eignung als Testitem stark ein.

Die unreflektierte Verwendung von Genres als Testitems, die Rentfrow und Gosling vorgeworfen wird, bezieht sich aber nicht nur auf ihren Entstehungsprozess innerhalb eines Individuums, sondern auch auf ihre gesellschaftliche Bedeutung und die heterogene Struktur innerhalb der Genres selbst. In seiner »Theory of Musical Genres« spricht Fabbri (1982) von Genres als »set of musical events whose course is governed by a definite set of socially accepted rules«, wobei er die »musical events« als »any type of activity performed around any type of event involving sound« definiert (ebd.: 1). Dabei lässt sich Bezug nehmen zum socialized nuclear content, den man mit den hier erwähnten sozial akzeptierten Regeln gleichsetzen kann. Diese Regeln unterteilt Fabbri in verschiedene Typen, wobei er auf die Unvollständigkeit der Liste hinweist und ebenfalls auf die Notwendigkeit einer interdisziplinären Forschung eingeht, um mehr dieser Regeltypen identifizieren zu können. Außerdem gibt er keine hierarchische Ordnung der Regeln vor, weist aber daraufhin, dass je nach Genre einige wichtiger als andere oder vollkommen irrelevant sein können, was durch die jeweilige Genre-Ideologie oder »Hyperrule« begründet ist, die diese Hierarchisierung bestimmen kann (ebd.: 2f.).³

3 Fabbri beschreibt zuerst die formalen und technischen Regeln. Diese umfassen unter anderem die typische Form, Komposition und die Struktur. Aber auch der Aufbau und Inhalt des Songtextes, falls vorhanden, genauso wie seine Verbindung zum Sound können zur Individualisierung eines Künstlers als auch eines Genres beitragen (Fabbri 1982: 3). Die semiotischen Regeln schaffen die Beziehung zwischen Inhalt und musikalischem Event. Dies können bestimmte verwendete Symboliken in Songtexten oder in Bildform sein, bestimmte Orte und

Unter Berücksichtigung dieser Genretheorie ist es zweifelhaft, Genrebegriffe als Testitems zu verwenden, wie es im STOMP geschieht. Dort wird eine Evaluation eines Begriffes verlangt, der wesentlich komplexer ist und mit verschiedenen Bewertungsgrundlagen unterschiedlich beurteilt werden kann. Eine Person könnte von der Musik eines Genres auf der formalen und technischen Ebene einer Metal-Performance sehr angetan sein, gleichzeitig aber dem üblichen Kleidungs- oder Tanzstil wie dem »Moshen« oder »Pogen« ablehnend gegenüberstehen. Solche Ambivalenzen können im STOMP nicht zum Ausdruck gebracht werden. Es ist außerdem nicht eindeutig, auf Basis welcher Regeln eine Person die Evaluation abgibt, wodurch die Vergleichbarkeit der Ergebnisse des STOMPs nicht möglich ist. Auch Halo-Effekte, bei denen ein Aspekt eines assoziierten Genres alle anderen Regeln dieses Genres überstrahlt, wie es auch bei Vorurteilen der Fall ist, können nicht kontrolliert werden, was ebenfalls die Validität des Testes einschränkt. Es könnten für den Probanden auch deshalb hauptsächlich bestimmte Regeln im Vordergrund stehen, weil er sich mit den anderen überhaupt nicht auskennt. Eine Person mit guten Kenntnissen im Bereich der formalen und technischen Regeln, beispielsweise ein Musiker, wird auf dieser Ebene vielleicht eine differenziertere Betrachtung vornehmen, während Nichtmusiker andere Regeln für die Bewertung des Genres in den Vordergrund stellen. Dies führt wiederum zu der Schlussfolgerung, dass der Test kein valides Maß für den Musikgeschmack oder die Musikpräferenzen sein kann, da es bei jedem Teilnehmer individuell zu einer Bewertung nach eigenen Kriterien kommt, die überhaupt nichts mit der klingenden Musik zu tun haben muss.⁴

Zeitpunkte des musikalischen Events, aber auch Kleidung, Instrumente, Tanz, Mimiken und bestimmte Kritiker. Ebenso spielt hier die Kommunikationsfunktion der Musik eine Rolle: Sie kann beispielsweise auf Emotionalität abzielen, auf Poesie oder Aggression (ebd.: 4). Verwandt mit den semiotischen Regeln sind die Verhaltensregeln, die sowohl Musiker, als auch Zuschauer/Fans umfassen und sich in bestimmten Reaktionsmustern, Ritualen und Verhalten gegenüber Außenstehenden äußern können. Sozialwissenschaftlich gesehen dürften besonders die sozialen und ideologischen Regeln interessant sein, die die soziale Struktur in der Genre-Community, ihre Hierarchien, Klassen, Gruppen und Generationen prägen sowie den Umgang mit weiteren Codes und Regeln beeinflussen. Abschließend dürfen auch die ökonomischen und rechtlichen Regeln nicht vergessen werden – auch wenn über sie häufig wenig bekannt gegeben wird –, da sie für die langfristige Existenz eines Genres mitverantwortlich sind (ebd.: 5).

- 4 Es wurde bereits angedeutet, dass Genres auch der Veränderung unterworfen sind, eine Tatsache, der Fabbri ebenfalls Rechnung trägt, wenn er über die Entstehung und das Vergehen von Genreregeln bzw. -codes spricht. Viele Regeln werden dabei von anderen Genresystemen von der jeweiligen Community adaptiert, wenn sie sich als erfolgreich erwiesen haben, oder aber selbst entwickelt

Interpretation der ermittelten Faktoren

Rentfrow und Gosling (2003c: 1242) konnten mit einer explorativen Faktorenanalyse vier Faktoren extrahieren, die von ihnen mit den Begriffen »Reflective and Complex«, »Intense and Rebellious«, »Upbeat and Conventional« und »Energetic and Rhythmic« genauer beschrieben wurden. Diese Benennungen resultieren aus den Genres, die auf die jeweiligen Faktoren laden. Auf Basis der hohen Faktorladungen von Blues, Jazz, Classical und Folk auf den ersten Faktor, bezeichneten Rentfrow und Gosling diesen mit dem Namen Reflective and Complex. Mit der entsprechenden Berücksichtigung der Genretheorie von Fabbri stellt sich aber die Frage, was nun genau reflektiv und komplex an diesen Genres ist und auf welcher Ebene die Autoren sich bei der Interpretation bewegt haben. Es ist nicht eindeutig, ob nun speziell die Musik gemeint ist, Songtexte, soziale Strukturen im Genre oder das Regelgefüge dieser Genres. Selbst wenn man annimmt, dass sich die Beschreibungen auf die musikalischen Strukturen beziehen, ist die Zuordnung von reflektiv und komplex für diesen Faktor eine Einigung des Forschungsteams auf Basis der eigenen Erfahrungen, der von den Einstellungen anderer zu diesen Genres abweichen kann. Somit sind diese Bezeichnungen von star-

– ein Prozess, den Fabbri (1982: 6f.) als »codification« bezeichnet und der durch zahlreiche Faktoren begründet wird, wie die technische Entwicklung oder die Verkündung ästhetischer Programme. Diese Codes werden aber nur solange akzeptiert, wie sie von der Mehrheit der Mitglieder als reichhaltig bzw. komplex genug angesehen werden und damit das Interesse bewahren können. Laut Fabbri sorgen gerade komplexe Regelsysteme für eine längere Stabilität eines Genres, da sie genug Unterhaltung bieten und damit keinen Wunsch nach Veränderung aufkommen lassen (ebd.: 7). Eine solche Behauptung lässt sich insofern empirisch untermauern, da Daniel Berlyne bereits 1971 einen umgekehrt u-förmigen Zusammenhang zwischen der Komplexität eines Stimulus und dem Gefallen entdeckte (vgl. Hargreaves/North 2011: 522). Demnach werden mittlere Komplexitäten gegenüber anderen bevorzugt. Allerdings wird von Fabbri die Einschränkung gemacht, dass die Wahrnehmung dieser Komplexität vom Kompetenzvermögen jedes Zuhörers, Kritikers und auch Künstlers abhängig ist und damit zu interindividuell unterschiedlichen Interpretationen von Codes, wie einer Ideologie, führen kann. Dies kann die unterschiedlichen Interpretationen von Codes/Regeln erklären und damit auch als Grundstein für die musikalische Vielfalt gelten (Fabbri 1982: 8). Darüber hinaus erinnert er daran, dass Erfolg und Interesse nicht für alle Genres Teil der Motivation für musikalische Aktivität sind – als Beispiel werden Genres angeführt, bei denen Musik Ritualcharakter besitzt – und dies zum einen kulturelle Unterschiede bei Genredefinitionen und -ausgestaltung verdeutlicht und zum anderen erklärt, wie Genreregeln sowohl den Gebrauch anderer Genreregeln als auch den Codierungsprozess und den Einfluss analytischer Kompetenz im Bezug auf diese Regeln beeinflussen (ebd.: 7).

ker Subjektivität geprägt. Ob die Interpretation der Faktoren, wie sie vom Forscherteam vorgenommen wurde, die richtige Erklärung ist, warum die entsprechenden Genres hauptsächlich auf einen Faktor laden, kann deswegen infrage gestellt werden. Es spielten neben ihrem Wissen auch die für Genres typischen Überzeugungen und Stereotype der Forscher eine große Rolle, über deren Sachkenntnis auf dem Gebiet der Musik sich nur spekulieren lässt. Dabei ist gerade ein gewisses musikalisches Verständnis und das damit verbundene Eintauchen in die komplexe soziale, musikalische und kulturelle Welt eines Genres notwendig, um solche Beschreibungen treffend fassen zu können. Dieses Verständnis fehlt Rentfrow und Gosling, weswegen ein Heranziehen von Experten notwendig gewesen wäre. Die Bezeichnungen der Faktoren, die zwischen den Psychologen ausgehandelt wurden, decken sich daher teilweise nicht mit den verwendeten Genreinhalten. Wenn man beispielsweise Soundtracks/theme songs betrachtet, lädt dieses Genre am höchsten auf den als Upbeat and Conventional bezeichneten Faktor (Rentfrow/Gosling 2003c: 1242, Tabelle 1). Da das Genre aber an sich aufgrund seiner Heterogenität die verschiedensten Musikrichtungen beinhalten kann, von Rockmusik bis Klassik, die laut Ergebnissen auf andere Faktoren laden, ist die Verwendung der Bezeichnung Upbeat and Conventional irreführend und stellt eine zu große Vereinfachung dar. Das Ergebnis ist genau genommen sogar ein Widerspruch zum Item. Das Genre lädt lediglich auf einen einzigen Faktor hoch, müsste aber nach der Benennung der Faktoren auch von den anderen Faktoren wesentlich höher determiniert sein, da rebellische Soundtracks genauso existieren wie auch sehr energiegeladene oder reflektive.

Gleiches gilt auch für das Item Rap/Hip-Hop. Dieses Genre kann sowohl politische als auch gewaltverherrlichende Thematiken aufgreifen (Wicke/Ziegenrucker/ Ziegenrucker 2007: 579f.), lädt aber stattdessen auf die Faktoren Intense and Rebellious und Reflective and Complex leicht negativ. Ähnliche Widersprüche finden sich auch bei ihrem Genre Religious. Wenn Musik, die als Ausdruck einer bestimmten Religion fungiert, lediglich als optimistisch und gebräuchlich bezeichnet wird, werden sämtliche religiöse und damit vor allem komplexe, reflektive Aspekte dieser Musik verneint, obwohl sich besonders diese Musikrichtung mit tiefgründigen Thematiken wie Glauben, Weltanschauungen und Traditionen auseinandersetzt.

Fazit

Betrachtet man die Ergebnisse der kritischen Auseinandersetzung, ergibt sich ein ernüchterndes Bild. Rentfrow und Gosling haben die bisherige Forschung im Bereich des Musikgeschmacks und der -präferenzen nicht ausreichend berücksichtigt und testtheoretische Standards bei der Erstellung des STOMP vernachlässigt. Sie haben ihre eigene Forschung unzureichend mit theoretischem Fundament untermauert bzw. diese Notwendigkeit durch die induktive Testkonstruktion umgangen und gleichzeitig wichtige Voraussetzungen für eine stabile Testkonstruktion, z.B. die Konstruktdefinition, nicht geschaffen. Stattdessen basieren sie ihre Studienreihe sehr stark auf Vermutungen und augenscheinlichen Zusammenhängen und gehen insgesamt zu unreflektiert an die Thematik heran. Des Weiteren unterschreiten Rentfrow und Gosling häufig gewisse Standards des wissenschaftlichen Arbeitens bei der Testerstellung wie bei der Datenerhebung oder ignorieren sie gänzlich.

Diese Mängel gehen einher mit einer zu oberflächlichen Arbeitsweise. Verwendete Literatur und Studien werden nicht vollständig berücksichtigt, Faktoren, die die Aussagekraft des STOMP einschränken, werden nicht erwähnt. Die nachträgliche Betrachtung des Genrebegriffs in Bezug auf seine Eignung als Testitem fällt ebenfalls negativ aus. Die Genretheorien und Überlegungen zur Bildung menschlicher Taxonomien zeigen, dass sich hinter dem Begriff eines Genres eine hochkomplexe und vielseitige Welt aus Überzeugungen, Codes und Strukturen verbirgt, die je nach Vorwissen und Einstellungen unterschiedlich wahrgenommen und beurteilt werden kann und keinesfalls ein stabiles Gefüge darstellt. Ohne die Einordnung in einen gewissen Kontext ist nicht ersichtlich, auf welcher Basis eine Evaluation des Genres vorgenommen wurde. Der STOMP erfasst folglich die Bewertung der subjektiven Repräsentation, die der Teilnehmer beim Lesen eines Genrebegriffs entwickelt, ohne dass diese für Außenstehende begreifbar gemacht werden könnte. Es zeigt sich, dass die mangelnde interdisziplinäre Betrachtung der Thematik der Forschung von Rentfrow und Gosling in dieser Studienreihe geschadet hat.

Für die zentrale Fragestellung bedeutet dies, dass der STOMP gleich in mehrfacher Hinsicht kein valides Testinstrument für den Musikgeschmack eines Menschen darstellen kann. Dafür fehlt zunächst das theoretische Konstrukt, auf das STOMP Bezug nehmen kann. Darüber hinaus ist seine Erstellung mit zahlreichen empirischen und psychologischen Mängeln behaftet, die ihn als unzuverlässiges Testverfahren identifizieren, dessen Validität in keiner Weise gesichert worden ist. Zusätzlich verwendet man im STOMP Items, die als zu abstrakt und allgemein bezeichnet werden müssen und die ohne Kontext fern jeder sozialen Wirklichkeit bewertet werden. Gerade der letzte Punkt macht deutlich, dass auch die mit dem STOMP vergleichbaren

Tests wie der MPQ und der MPS und die mit ihnen durchgeführten Studien sich mit derselben Kritik konfrontiert sehen müssen (Cattell/Anderson 1954; Delsing et al. 2008; Dollinger 1993; Litle/Zuckerman 1986; Rawlings/Ciancarelli 1997; Selfhout et al. 2007). Damit haben ihre Ergebnisse im Grunde genommen keinerlei Aussagekraft im Bezug auf den Musikgeschmack bzw. die -präferenzen eines Individuums, genauso wenig wie nachfolgende Studien, in denen der STOMP verwendet wird und Korrelationen zwischen seinen Werten und anderen Persönlichkeitsmerkmalen berechnet werden (Sigg 2009; Zweigenhaft 2008). Um Musikpräferenzen zu erfassen, sind deutlich komplexere und ausführlichere Testverfahren notwendig, wobei sich die Frage stellt, wie ein solches Verfahren aussehen und auf welches theoretische Fundament es sich stützen könnte.

Ein neuer Ansatz für die Musikpräferenzmessung

Wenn ein Testverfahren erstellt werden soll, das wirklich einen Bezug zum Musikgeschmack einer Person hat, muss zunächst eine Theorie des Musikgeschmacks zugrunde gelegt werden, welche Musikpräferenz und Musikgeschmack eindeutig als Konstrukte definiert. Aus den bereits zu Beginn der Arbeit erwähnten Theorieansätzen eignet sich in musikpsychologischer Hinsicht besonders der Ansatz von Ekkehard Jost (1982), der Präferenzen mit Einstellungen gleichsetzt. Der Musikwissenschaftler sieht den Geschmacksbegriff wegen seiner Mehrdeutigkeit und ideologischen Besetzung verstärkt durch den Präferenzbegriff ersetzt, den er im Sinne der Sozialpsychologie den Einstellungen zurechnet (Jost 1982: 246). Für Jost sind sie durch Erfahrung erworbene Wertorientierungen, die durch den Sozialisationsprozess mitbestimmt werden und Erwartungen und Werthaltungen der soziokulturellen Umwelt reflektieren, in der sie erworben worden sind. Jost kennzeichnet die Einstellung im Gegensatz zu den temporären Dispositionen als relativ beständig und unabhängig von psychophysischen Zuständen oder situativen Bedingungen, sieht sie dennoch durch langfristige Einflüsse oder Schlüsselerlebnisse veränderbar. Darüber hinaus erkennt Jost in den Präferenzen den Steuerungsmechanismus musikalischen Verhaltens und den Bezugsrahmen für Urteile und Vorurteile. Außerdem seien diese mitverantwortlich für die Selektion musikalischer Wahrnehmungsinhalte und Kontinuität des Verhaltens und fungierten als Mittel sozialer Anpassung und Abgrenzung (ebd.: 247).⁵

Man könnte auf der Basis von Einstellungen quantitativ zwischen Musikgeschmack und -präferenz abstufen: In Anlehnung an Behnes Musikkonzept ließe sich der Musikgeschmack als das gesamte Bündel von Einstellungen gegenüber jeder Art von Musik und den mit ihr zusammenhängenden Aspekten betrachten, von dem man lediglich bestimmte Teilbereiche des Gesamtmerkmals, also einzelne Präferenzen, genauer betrachtet. Es würde den

5 Diese Entscheidung wird auch von Andreas C. Lehmann (1994: 112) geteilt, wenn er Präferenzen wie Einstellungen eine evaluative Komponente zuordnet, und beide sowohl verhaltensintentional auf einen Einstellungsgegenstand gerichtet seien als auch als durch die gezeigte Präferenz funktional erklärbar würden. Eine solche Übertragung der Musikpräferenz in die Einstellungsforschung hat den großen Vorteil, dass man ein bereits gut erforschtes Konstrukt verwenden kann, für das es bereits zahlreiche verschiedene Theorien und Konzepte gibt (Dawes 1977: 44ff.). Auch im Bereich der Testtheorie lassen sich mit den Einstellungstests eine Reihe von Skalierungsverfahren und Messtechniken isolieren, die als Basis für die Präferenzforschung verwendet werden können (ebd.: 9ff.; Petermann 1980: 22ff.).

Umfang dieses Artikels übersteigen, tiefer in die Einstellungskonzepte und Theorien einzusteigen. Stattdessen soll eine Theorie hervorgehoben werden, die speziell auf Musikpräferenzen zugeschnitten ist. Lehmanns Situation-Funktion-Präferenz-Modell (SFP-Modell) ist ein »kognitives Lernmodell zur funktional bedingten Präferenzentstehung und -änderung« (Lehmann 1994: 122). Dieser Ansatz ist deshalb so interessant, weil das Modell mehrere Einstellungstheorien in Bezug auf Präferenzen zusammenführt, besonders Behnes Hörstrategien als Ursache für Präferenzen, Icek Ajzen und Martin Fishbeins Theorie des geplanten Verhaltens und Harry C. Triandis Idee, dass Verhalten auch zukünftige Einstellungen beeinflusst (ebd.: 122ff.).⁶

Dieses Modell bezieht sich auf das konkrete Musikhören und unterscheidet sich bereits dadurch von den im STOMP erhobenen verbalen Präferenzen gegenüber einem Genre. Die Bestätigung des Modells könnte daher auch die Erklärung bieten, warum sich in der bisherigen Forschung verbale und klingende Präferenzen nicht als deckungsgleich erwiesen haben (Behne 1986: 26; Gebesmair 2001: 92f.; Gembris 2005: 284; Jost 1982: 260). Wenn beim Musikhören das Erleben der Musik evaluiert wird und beim Genrebegriff eine Vielzahl an Genre-codes einer Gesamtbewertung unterzogen werden, können nicht die gleichen Ergebnisse zustande kommen. Sollen beide Bereiche des Musikgeschmacks erfasst werden, müsste eine Kombination von verbalen und klingenden Präferenzen angestrebt werden.

6 Die aktuelle Situation mit ihren zeitlichen, räumlichen und personalen Konstanten wird vom Hörer kategorisiert. In Verbindung mit den aktuellen Bedürfnissen des Hörers wird anschließend die Funktion der Musik oder ihre intendierte Wirkung abgeleitet, falls ein Musikwunsch vorhanden ist (Lehmann 1994: 123). Eine bekannte Situation-Funktion-Kombination führt zu einem Rückgriff auf bereits bekannte Verhaltensmuster aus früheren Rezeptionssituationen. Eine positive Bewertung von aktuellem und vergangenem Verhalten kann zu einer zunehmenden Gewohnheitsbildung bei der Wahl der Präferenz- und Verhaltensmuster führen, wobei auch die Einstellung zu diesem Verhalten von Bedeutung ist. Bei unbekannten Konstellationen können mithilfe kognitiver Prozesse Vermutungen über Wirkungen und Funktionen angestellt werden (ebd.: 124). Nicht bekannte Situation-Funktion-Präferenz-Konstellationen werden in den Erfahrungsschatz integriert. Wichtig ist, dass während der Rezeption nicht die Musik, sondern das Erleben der Musik insofern bewertet wird, ob ein Hörerlebnis die erwartete Wirkung hat. Dies ist dem Hörer nicht bewusst, da Musik und Erleben eng miteinander gekoppelt sind. Bestimmte Musik wird also mit spezifischer Wirkung verbunden und eine zufriedenstellende Wirkung bedeutet wiederum eine zufriedenstellende Musik. Die Handlungsfolge hat dabei Rückwirkung auf die kognitiven Prozesse wie Überzeugungen und die bewertende Komponente einer Einstellung. Dies kann man als lerntheoretischen Aspekt interpretieren, der auch langfristige Änderungen von Präferenzen durch Hemmung und Verstärkung ermöglicht, da die Wissensgrundlage verändert wird: »Die Funktion wird dabei zur intervenierenden Variablen zwischen den Antezedenzbedingungen und der gezeigten Präferenz« (ebd.: 124).

Kanones innerhalb von Genres als Einstellungsgegenstand

Schwieriger ist trotz vorhandener Einstellungsmodelle die Festlegung der Einstellungsobjekte. Genres haben sich zu sehr als instabil erwiesen, als dass sie als Items zu verwenden wären. Es wäre dennoch falsch, sie gänzlich zu verwerfen. Trotz ihrer fluiden Eigenschaften sind diese Kategorien für Fabian Holt (2007: 180) verantwortlich für Struktur und Abgrenzung innerhalb der Musikkultur und spiegeln ihre Komplexität wider, sodass eine Untersuchung von Genres und ihrer Verzweigungen helfen kann, das soziale Leben besser zu verstehen.

Um ein langfristig stabiles Testverfahren zu schaffen, müssten Teilbereiche von Genres identifiziert werden, die über einen längeren Zeitraum kaum Veränderungen unterworfen sind. Einen solchen stabilen Teil könnten Kanones darstellen.⁷ Diese können mehrere Funktionen erfüllen. Zum einen bieten sie die Möglichkeit der Informationsreduktion und Fokussierung auf die wichtigsten Aspekte einer Gesellschaft oder Kultur und erfüllen damit Orientierungs- und Sicherheitsfunktionen. Zum anderen werden die im Kanon gesammelten Objekte historisiert und für spätere Generationen als kulturelles Gedächtnis bewahrt. Dadurch können sie genauso als Medium der Sozialisation verwendet werden wie als Reflexionspunkt oder auch als Herrschafts- oder Disziplinierungsinstrument (Anz 1998: 6). Ein akzeptierter Kanon übt also immer auch eine Macht und Ideologie aus, da er als Maßstab für Bewertungen herangezogen werden kann. Außerdem kann er als eine Art aktiver Normsender fungieren, indem er einen anzustrebenden Idealzustand repräsentiert, und als Legitimation geltender Werte gesehen werden (Heydebrand 1998: 617). Voraussetzung für ihn ist aber, dass eine bestimmte Gruppe die Wertvorstellungen im Bezug auf diesen Kanon teilt, den Kanon anerkennt und seine Bewahrung, Pflege und Stärkung unterstützt. Der Kanon hilft ihnen bei der individuellen und kollektiven Identitäts- und Sinnstiftung und Abgrenzung zu anderen Gruppen. Daher übt der Kanon seine

7 Im Allgemeinen bedeutet der aus dem griechischen stammende Begriff Kanon »Maßstab« oder auch »festgesetzte Ordnung« (Appen/Doehring/Rösing 2008: 25). Er beschreibt eine Sammlung von Autoren, Werken, Riten, Verhaltenstypen und Stilen, aber auch Kriterien und Methoden, die als Ergebnis von Selektionsentscheidungen und damit verbundenen Wertungsakten als so wertvoll erachtet wurden, dass sie tradiert werden (Heydebrand 1998: 613). Kanones sind demzufolge Produkte sozialer Handlungen in sozialen Gruppen und damit auch (Sub-)Kulturen und Genres. Je nach Inhalt und verwendeten Strukturmodellen lassen sich verschiedene Arten von Kanones differenzieren.

normative Kraft meist nur in diesen speziellen Subsystemen aus, kann aber durchaus ins öffentliche Bewusstsein sedimentieren (Appen/Doehring/Rösing 2008: 45f.).

Ein solches System stellt auch ein Genre dar, in dem Künstler, Musikindustrie, Kritiker und Fans, verbunden durch Erwartungen, Überzeugungen und Konventionen, für eine bestimmte Art von Musik mitverantwortlich sind (Lena/Peterson 2008: 698). Zu diesem Schluss kommt auch Holt (2007: 17): »Genre formation is also canon formation«. Gestaltung und Grenzen des Genres sind zwar beständiger Teil eines gemeinsamen Aushandelns dieser Werte, in einigen Szenen lässt sich aber bereits eine Historisierung ihrer Werte feststellen. Jennifer C. Lena und Richard A. Peterson (2008: 706) entdeckten unterschiedliche Verlaufsformen der Genreentwicklung und identifizierten die eine Entwicklung abschließende »traditionalist-phase«, in der mit der Kanonisierung begonnen wird und sich Standards etablieren. Dies können Performances, Techniken, Rituale, Künstler und musikalische Werke sein, womit sich im Bezug auf Fabbris Genretheorie wahrscheinlich Kanones für alle Ebenen der Genreregeln einer Musikrichtung finden lassen, die als die Kernwerte und Ideale des Genres angesehen werden können.

Kanones sind zeitlich stabil und so für die Präferenzforschung als Einstellungsobjekte geeigneter als andere vergängliche Teilbereiche eines Genres. Man könnte die exemplarischen Musikstücke auf Basis des SFP-Modells genauer betrachten, sofern die Theorie bestätigt werden kann, und die Einstellungen zu den anderen kanonisierten Genrecodes mithilfe von verbalen Einstellungstests erfassen. So ließe sich ein sehr differenziertes Bild über die Einstellungen zu Kerninhalten eines Genres zeichnen, das auch mögliche Ambivalenzen darstellen kann und Musik samt sozialem Kontext des Genres gleichzeitig betrachtet. Als Voraussetzung dafür müsste man aber einen großen Erhebungsaufwand für die Erfassung der Kanones in Kauf nehmen und könnte nur Genres untersuchen, in denen bereits Kanonisierungsprozesse stattgefunden haben. Außerdem müssten die Kanonstrukturen im Genre kritisch beachtet werden, da es konkurrierende Kanones oder bewusste Abgrenzungen von einem Hauptkanon geben kann, die eventuell auf Abspaltung und Entstehung neuer Genres und Stile hinweisen können. Dadurch hätte das Testverfahren wie schon der STOMP nie Anspruch auf eine vollständige Erfassung aller Genres. Ebenfalls müssten regionale Unterschiede berücksichtigt werden, die sowohl bestimmte Genres länderspezifisch machen würden als auch innerhalb eines Genres stabile Unterschiede hervorrufen könnten. Ein langfristig stabiler Musikgeschmack muss sich in den Testpersonen ebenfalls etabliert haben. Daher wäre ein Testverfahren erst für Personen in der Abschwungphase sinnvoll. Gleichzeitig muss die

Stabilität der Kanones regelmäßig überprüft werden, damit der Test an eventuelle Schwankungen angepasst werden kann. Dies hätte sehr zeit- und kostenaufwendige Normierungen zur Folge.

Die dabei entwickelte Itemmenge für ein Genre allein wäre wahrscheinlich groß genug, um einen eigenen Test für jedes Genre zu gestalten. Es müssten sowohl ethnographische Beobachtungen als auch Interviews mit Musikern und Fans durchgeführt werden, um semiotische, soziale und Verhaltensregeln zu identifizieren. Für die ökonomischen Regeln wären Kooperationen mit Teilen der Musikindustrie notwendig. Musikzeitschriften und Kritiker stellen dagegen häufig öffentlich Ranglisten der ihrer Meinung nach besten Songs eines Genres zusammen, die man kombinieren könnte, um sie als Quelle für exemplarische Werke zu verwenden (Appen/Doehring/Rösing 2008: 34). Da dieses Verfahren große Ähnlichkeit mit der von Rentfrow und Gosling verwendeten Methode der fünften Studie, Songs von »Essential Compilations« als Prototypen zu verwenden, aufweist, birgt es auch die gleichen Risiken (Rentfrow/Gosling 2003c: 1245). Von den Präferenzen für die kanonisierten Songs kann nicht direkt auf die Präferenzen für aktuelle Songs des Genres geschlossen werden, da der Zusammenhang zwischen ihnen unklar ist. Die kanonisierten, meist wesentlich älteren Songs können die Basis für die Entwicklung späterer Songs liefern, müssen aber nichts mit ihnen gemein haben oder gleich präferiert werden. Denn innerhalb eines Genres lassen sich zahlreiche Subgenres oder Stile identifizieren, die sich später entwickelten und zu denen der Zugang leichter fallen kann. Dennoch könnte die Kanonisierung eine solche Macht auf die Entwicklung des Genres gehabt haben, dass bestimmte Elemente sich durch diese Songs ziehen. Es wäre ein besonders wichtiger Teil des Verfahrens, die Vorhersagekraft von Präferenzen für prototypische, kanonisierte Werke eines Genres für Präferenzen außerhalb des Kanons zu ermitteln. Gleichzeitig muss beachtet werden, dass das SFP-Modell immer Musik in einen situativen Kontext setzt und die Bedürfnisse in dieser Situation berücksichtigt (Lehmann 1994: 123). Die im Test hergestellte Situation wäre sehr konstruiert, sodass hier die externe Validität, also die Übertragbarkeit der Situation in den Alltag, problematisch wäre. Gleichzeitig müsste deshalb nachgefragt werden, in welchen Situationen die Musik präferiert werden würde, was wiederum eine Bewertung imaginierter Situationen verlangt, die ebenfalls von der Realität abweichen können (Schramm/Vorderer 2002: 114).

Letztendlich muss besonders berücksichtigt werden, dass man sich nah an der Grenze zu Stereotypen gegenüber verschiedenen Musikrichtungen bewegt. Es ist besonders viel Aufmerksamkeit und Vorsicht geboten, damit ein solches Testverfahren weder Vorurteile vermittelt noch die Genres zu

stark auf die kanonisierten Werte reduziert, da es generell nur die Kernorientierungen in den Präferenzen erfassen könnte. Musikalische Vielfalt entsteht allerdings erst dann, wenn man von bekannten Normen abweicht und neue Sounds bzw. Stile abseits der etablierten sucht.

Ausblick

Es ist schwierig zu beurteilen, ob eine solch aufwendige Erstellung eines standardisierten quantitativen Messverfahrens für Musikpräferenzen lohnenswerter wäre als ein einstündiges qualitatives Interview, in dem man sich ebenfalls ein komplexes Bild von den Prä- und Postferenzen einer Person machen könnte. Für die Musikindustrie wäre beides keine sinnvolle Erweiterung ihrer Marktforschungsinstrumente, da man sich für das strategische Marketing eher an Chartplatzierungen und Verkaufszahlen orientiert, um Trends zu generieren oder zu verfolgen, anstatt individuelle Präferenzen zu analysieren. Der Mehrwert eines quantitativen Verfahrens läge besonders in der Einbindung des sozialen Kontextes eines Genres, der auch Aufschluss über die Zugehörigkeit zu Szenen und Subkulturen geben kann. Das Verfahren könnte die Verinnerlichung der dazugehörigen Normen und Vorstellungen messbar machen, wodurch es besonders für die Sozialwissenschaften interessant wird. Wenn Zusammenhänge mit anderen Merkmalen wie der Persönlichkeit genauer untersucht werden sollten, wie Rentfrow und Gosling es in ihrer Studienreihe intendierten, wäre ein solch quantitatives Testverfahren ebenfalls unabdingbar. Es zeigt sich, dass es beim Musikgeschmack um viel mehr geht als nur den konkreten musikalischen Stimulus. Musik als ein Kommunikationsmedium übermittelt auch Werte, Überzeugungen und Symbole des jeweiligen Genres; diese gehen in die Präferenzen mit ein. Es ist daher notwendig, diese Multidimensionalität in der Forschung stärker zu berücksichtigen, um auch der Komplexität von Kultur gerecht zu werden. Mit den hier vorgeschlagenen Ansätzen ist ein Schritt in diese Richtung getan.

Literatur

- American Psychological Association (2010). *Publication manual of the American Psychological Association*. Washington, DC: American Psychological Association (6. Aufl.).
- Anz, Thomas (1998). »Einführung.« In: *Kanon – Macht – Kultur. Theoretische, historische und soziale Aspekte ästhetischer Kanonbildungen*. Hg. v. Renate von Heydebrand (= Germanistische Symposien Berichtsbände 19). Stuttgart: Metzler, S. 3-8.
- Appen, Ralf von / Doehring, André / Rösing, Helmut (2008). »Pop zwischen Historismus und Geschichtslosigkeit: Kanonbildungen in der populären Musik.« In: *No time for losers. Charts, Listen und andere Kanonisierungen in der populären Musik*. Hg. v. Dietrich Helms und Thomas Phleps (= Beiträge zur Populärmusikforschung 36). Bielefeld: transcript, S. 25-50.
- Behne, Klaus-Ernst (1986). *Hörertypologien: Zur Psychologie des jugendlichen Musikgeschmacks*. Regensburg: Bosse.
- Bühner, Markus (2006). *Einführung in die Test- und Fragebogenkonstruktion*. München: Pearson Studium (3. akt. u. erw. Aufl.).
- Cattell, Raymond B. / Anderson, Jean C. (1953). »Ipat music preference test of personality.« In: *Handbook for the IPAT Music Preference Test of Personality*, <http://web.ebscohost.com/ehost/detail?vid=3&hid=25&sid=142bf6c6-39e7-4161-b20b8608716655ef%40sessionmgr14&bdata=Jmxhbm9ZGUmc2l0ZT1laG9zdC1saXZl&db=pdx&AN=PT9001294>, Zugriff: 30.5.2012.
- Chamorro-Premuzic, Tomas / Swami, Viren / Furnham, Adrian / Maakip, Ismail (2009). »The big five personality traits and uses of music« In: *Journal of Individual Differences* 30, Nr. 1, S. 20-27.
- Dawes, Robyn M. (1977). *Grundlagen der Einstellungsmessung*. Weinheim: Beltz.
- Delsing, Marc J. M. H. / ter Bogt, Tom F. M. / Engels, Rutger C. M. E. / Meeus, Wim H. J. (2008). »Adolescents' music preferences and personality characteristics.« In: *European Journal of Personality* 22, Nr. 2, S. 109-130.
- Deutsche Eliteakademie (2011) »Veröffentlichung von Studien und die damit verbundene Publikationsethik.« In: <http://www.eliteakademie.biz/load.php?name=Print&art=Artikel&sid=250>, Zugriff: 30.5.2012 (Stand vom 4.7.2011).
- Dollase, Rainer (1998). »Musikpräferenzen und Musikgeschmack Jugendlicher.« In: *Handbuch Jugend und Musik*. Hg. v. Dieter Baacke. Opladen: Leske + Budrich. S. 341-368.
- Dollinger, Stephen J. (1993). »Research note: Personality and music preference: Extraversion and excitement seeking or openness to experience?« In: *Psychology of Music* 21, Nr. 1, S. 73-77.
- Fabbri, Franco (1982). »A theory of musical genres: Two applications.« In: *Popular Music Perspectives*. Hg. v. David Horn, Philip Tagg und Marian Green. Göteborg: International Association for the Study of Popular Music, S. 52-81.
- Fabbri, Franco (1999). »Browsing Music Spaces: Categories and the Musical Mind.« In: <http://www.tagg.org/others/ffabbri9907.html>, Zugriff: 30.5.2012.
- Gebesmair, Andreas (2001). *Grundzüge einer Soziologie des Musikgeschmacks*. Wiesbaden: Westdeutscher Verlag.
- Gembris, Heiner (2005). »Musikalische Präferenzen.« In: *Spezielle Musikpsychologie*. Hg. v. Rolf Oerter und Thomas H. Stoffer (= Enzyklopädie der Psychologie D VII 2). Göttingen: Hogrefe, S. 277-342.

- Hargreaves, David. J. / North, Adrian C. (2011). »Experimental aesthetics and liking for music.« In: *Series in affective science. Handbook of music and emotion. Theory, research, applications*. Hg. v. Patrick N. Juslin und John A. Sloboda. Oxford: Oxford University Press, S. 515-546.
- Heydebrand, Renate von (1998). »Kanon Macht Kultur: Versuch einer Zusammenfassung.« In: *Kanon – Macht – Kultur. Theoretische, historische und soziale Aspekte ästhetischer Kanonbildungen*. Hg. v. ders. (= Germanistische Symposien Berichtsbände 19). Stuttgart: Metzler, S. 612-625.
- Holt, Fabian (2007). *Genre in Popular Music*. Chicago: University of Chicago Press.
- John, Oliver P. / Hampson, Sarah E. / Goldberg, Lewis R. (1991). »The basic level in personality-trait hierarchies - studies of trait use and accesability in different contexts.« In: *Journal of Personality and Psychology* 60, Nr. 3, S. 348-361.
- Jost, Ekkehard (1982). »Sozialpsychologische Dimensionen des musikalischen Geschmacks.« In: *Systematische Musikwissenschaft*. Hg. v. Carl Dahlhaus und Helga de La Motte-Haber (= Neues Handbuch der Musikwissenschaft 10). Wiesbaden: Athenaion, S. 245-268.
- Lehmann, Andreas C. (1994). *Habituelle und situative Rezeptionsweisen beim Musikhören. Eine einstellungstheoretische Untersuchung*. (= Schriften zur Musikpsychologie und Musikästhetik 6). Frankfurt am Main: Lang.
- Lena, Jennifer C. / Peterson, Richard A. (2008). »Classification as culture: Types and trajectories of music genres.« In: *American Sociological Review* 73, Nr. 5, S. 697-718.
- Lienert, Gustav A. / Raatz, Ulrich (1998). *Testaufbau und Testanalyse*. Weinheim: Beltz (6. Auflage).
- Litle, Patrick / Zuckerman, Marvin (1986). »Sensation seeking and music preferences.« In: *Personality and Individual Differences* 7, Nr. 4, S. 575-578.
- Mulder, Juul / Ter Bogt, Tom F. M. / Raaijmakers, Quinten A. / Nic Gabhainn, Saoirse / Sikkema, Paul (2010). »From Death Metal to R&B? Consistency of music preferences among Dutch adolescents and young adults.« In: *Psychology of Music* 38, Nr. 1, S. 67-83.
- Niketta, Reiner (1993). »Musikpräferenzen und Musikgeschmack.« In: *Musikpsychologie. Ein Handbuch*. Hg. v. Herbert Bruhn, Rolf Oerter und Helmut Rösing. Hamburg: Rowohlt, S. 339-354.
- Petermann, Franz (Hg.) (1980). »Einstellungsmessung und -forschung: Grundlagen, Ansätze und Probleme.« In: *Einstellungsmessung, Einstellungsforschung*. Göttingen: Hogrefe, S. 9-36.
- Rawlings, David / Ciancarelli, Vera (1997). »Music preference and the five-factor model of the neo personality inventory.« In: *Psychology of Music* 25, Nr. 2, S. 120-132.
- Rentfrow, Peter J. / Gosling, Samuel D. (2003a). *Normative Data for the Four Music Preference Dimensions of the Short Test Of Music Preferences (STOMP)*. Unveröffentlichte Daten. In: <http://homepage.psy.utexas.edu/homepage/faculty/gosling/tipi%20site/STOMP%20norms.pdf>, Zugriff: 30.5.2012.
- Rentfrow, Peter J. / Gosling, Samuel D. (2003b). *STOMP [Verbaler Präferenztest]*. In: <http://homepage.psy.utexas.edu/homepage/faculty/gosling/tipi%20site/stomp.pdf>, Zugriff: 30.5.2012.
- Rentfrow, Peter J. / Gosling, Samuel D. (2003c). »The do re mi's of everyday life: The structure and personality correlates of music preferences.« In: *Journal of Personality and Social Psychology* 84, Nr. 6, S. 1236-1256.

- Rentfrow, Peter J. / Gosling, Samuel D. (2011). *STOMP-Revised* [Verbaler Präferenztest]. In: http://homepage.psy.utexas.edu/homepage/faculty/gosling/tipi%20site/STOMP_files/STOMPR.doc, Zugriff: 30.5.2012.
- Schramm, Holger / Vorderer, Peter (2002). »Musikpräferenzen im Alltag: Ein Vergleich zwischen Jugendlichen und Erwachsenen.« In: *Jugendforschung. Wozu Jugendliche Musik und Medien gebrauchen. Jugendliche Identität und musikalische und mediale Geschmacksbildung*. Hg. v. Renate Müller, Patrick Glogner, Stefanie Rhein und Jens Heim. Weinheim: Juventa, S. 112-125.
- Selfhout, Maarten H. W. / Delsing, Marc J. M. H. / ter Bogt, Tom F. M. / Meeus, Wim H. J. (2007). »Heavy metal and hip-hop style preferences and externalizing problem behavior: A Two-Wave Longitudinal Study.« In: *Youth & Society*, 39, Nr. 4, S. 435-452.
- Sigg, N. (2009). *An investigation between music preferences, personality and psychological wellbeing* In: <http://aut.researchgateway.ac.nz/bitstream/handle/10292/955/SiggN.pdf?sequence=4>, Zugriff: 30.05.2012.
- Wicke, Peter / Ziegenrucker, Kai-Erik / Ziegenrucker, Wieland (2007). *Handbuch der populären Musik: Geschichte, Stile, Praxis, Industrie*. Mainz: Schott (erw. Neuausg.).
- Zweigenhaft, Richard L. (2008). »A do re mi encore.« In: *Journal of Individual Differences* 29, Nr. 1, S. 45-55.

Abstract

The present article examines the study series »The Do Re Mi's of Everyday Life: The Structure and Personality Correlates of Music Preferences,« carried out by Rentfrow and Gosling in 2003, and focuses on the creation of the »Short Test of Music Preferences« and the registration of the preference dimensions in the second study. Thus, in interdisciplinary research, the previous research of music taste, genre and test theory are used as the standard of comparison. Numerous test-theoretical deficiencies, missing compliances of psychological quality criteria and insufficient definitions of concept are revealed. Genre terms prove to be too abstract and thus cannot capture the complexity of underlying social and cultural systems. Ambiguous and inappropriate terms are used for the interpretation of the factors without considering music experts or music taste research. Subsequently, it is recommended to equate music preferences with the established attitude construct. In addition, stable, canonised genre codes should be used as preference guidelines.